

**ABSTRACT**

A method and apparatus for composing and presenting multimedia programs using the MPEG-4 standard at a multimedia terminal (100). A composition engine (120) maintains and updates a scene graph (124) of the current objects, including their relative position in a scene and their characteristics, and provides a corresponding list of objects (126) to be displayed to a presentation engine (150). In response, the presentation engine begins to retrieve the corresponding decoded object data that is stored in respective composition buffers (176, . . . 186). The presentation engine assembles the decoded objects to provide a scene for presentation on output devices such as a video monitor (240) and speakers (242), or for storage. A terminal manager (110) receives user commands and causes the composition engine to update the scene graph and list of objects accordingly. The terminal manager also forwards the information contained in the object descriptors to a scene decoder (122) at the composition engine. Preferably, the composition and the presentation of the content are controlled using separate control threads to allow the presentation engine to retrieve and process the decoded object data while the composition engine is recovering additional scene description information and/or object descriptors.